Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:

1) **Mohd Danish:** **Email:** mdanish63364@gmail.com
   1) Feature Engineering:
      - Introduced new features
      - Introducing Dummy Variables
   2) Data Visualization:
      - Trend of Close price
      - Distribution of Close price
      - Heatmap
   3) VIF
   4) Regression Analysis:
      - Linear Regression
      - Lasso
      - Ridge
      - ElasticNet
   5) Group Colab
2) **Abdul Rahman Talha:** **Email:** rahman88talha@gmail.com
   1) Correlation Analysis
      - Between Independent Variables
      - Between Dependent and Independent Variables
   2) Data Visualization:
      - JointPlot
      - Distplot
   3) Regression Analysis:
      - Linear Regression
      - Lasso
      - Ridge
   4) PPT
3) **Huzaifa Khan:** **Email:** huzaifakhan2974@gmail.com
   1) Data Munging:
      - Introducing New variables
      - Settling DateTime
   2) Data Visualization:
      - Distplot
      - Boxplot
      - Scatter Plot
   3) Regression Analysis:
      - Linear Regression
      - Lasso
      - Ridge
      - ElasticNet

4) **Arbaz Malik:** **Email:** malikarbaaz267@gmail.com

1) Data Munging:
   - Feature Engineering
2) Data Visualization
   - Barplot
   - Scatter-Plot
3) Regression Analysis
   - Linear Regression
   - Lasso
   - Ridge
   - ElasticNet

Please paste the GitHub Repo link.

Github Link:-
https://github.com/Rahman88talha/Capstone_Project_Stock_Closing_Price_Prediction

Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

Yes Bank is a well-known bank in the Indian financial domain. Since 2018, it has been in the news because of the fraud case involving Rana Kapoor. Owing to this fact, it was interesting to see how that impacted the stock prices of the company and whether Time series models or any other predictive models can do justice to such situations. This dataset has monthly stock prices of the bank since its inception and includes closing, starting, highest, and lowest stock prices of every month. The main objective is to predict the stock's closing price of the month.

Our First Step was to import dataset through Pandas 'read_csv' then data wrangling and feature engineering in our dataset. We did not get into the situation to remove NA values because there is 0 null values in Yes Bank dataset.

Next, EDA(Exploratory Data Analysis) in which trend of stock closing price, distribution of dependent variable have been examined. Plotted histogram of all variables. Then log transformation has been applied on each variable, it leaded to a conclusion: to normalize right skewed data perform log transformation.

Now, correlation has been checked among each other through heatmap, there was very high correlation among independent features means high multicollinearity in our model, so to check how high multicollinearity is VIF(Variation Inflation Factor) has been checked on the basis of VIF.

Prepared independent and dependent variables for train test split method.

Applied Linear model, Ridge regression, Lasso regression and ElasticNet all the models are performing in a better way but Linear Model and Lasso is performing in better way in comparison to Ridge and ElasticNet but Ridge regression's and ElasticNet's performance improved by applying cross-validation and Hyperparameter tuning.

Conclusions:

1) Target Variable is strongly dependent on Independent Variables.
2) R2 is very high for every Regression which tells us that How our Independent variable describe dependent variable in dataset.
3) Predicted Model has shown similarity with actual Model .